As specifically shown in the examples, genes having the native nucleotide sequence can be obtained by, for example, screening cDNA libraries. DNA encoding a protein having a modified amino acid sequence can also be synthesized based on the DNA having the native nucleotide sequence by conventionally used site-directed mutagenesis or a PCR method. For example, a DNA fragment to be modified may be obtained by treating the native cDNA or genomic DNA with restriction enzymes, and using this as a template, site-directed mutagenesis or a PCR method is carried out using a primer into which the desired mutation has been introduced so as to obtain a DNA fragment into which the desired modification has been introduced. Then the mutation-introduced DNA fragment may be linked to a DNA fragment encoding another part of the protein of interest.

Alternatively, in order to obtain a DNA encoding a protein comprising a shortened amino acid sequence, a DNA encoding an amino acid sequence longer than the amino acid sequence of interest, for example a full-length amino acid sequence, is cleaved by a desired restriction enzyme, and when the resulting DNA fragment was found not to encode the entire amino acid sequence of interest, a DNA fragment comprising the lacking sequence is synthesized and ligated thereto.

By expressing the obtained gene using a gene expression system in *Escherichia coli* or yeast, the gene product MSH protein may be obtained. Alternatively, the MSH protein may be obtained by using an antibody against a protein encoded by the amino acid sequence as set forth in SEQ ID NO: 2 or 4. By using an antibody, the gene of a protein having a similar function to MSH may be cloned from another organism.

Thus, the present invention also relates to a recombinant vector comprising the above-mentioned gene, specifically an expression vector, and a host transformed with said vector. As a host, there can be used a

prokaryotic or eukaryotic organism. As a prokaryotic organism, there can be used such a common host as a microorganism belonging to the genus *Escherichia* such as *Escherichia coli*, a microorganism belonging to the genus *Bacillus* such as *Bacillus subtilis*, and the like.

As an eukaryotic host, there can be used a lower eukaryotic organism, for example an eukaryotic microorganism, for example a fungus, yeast or a mold. As yeast, there can be mentioned a microorganism belonging to the genus *Saccharomyces* such as *Saccharomyces cerevisiae*, and as a mold, there can be mentioned a microorganism belonging to the genus *Aspergillus* such as *Aspergillus oryzae* and *Aspergillus niger*, and a microorganism belonging to the genus *Penicillium*. Furthermore, animal cells or plant cells can be used: as animal cells, there can be used cell lines derived from mouse, hamster, monkey, human and the like, specifically COS cells, Vero cells, CHO cells, L cells, C127 cells, BALB/c3T3 cells, Sp-2/0 cells, and the like. As plant cells, there can be used cultured cells from tabacco, genus *Populus*, genus *Eucalyptus*, genus *Acacia*, and the like.

Insect cells such as silkworm cells or adult silkworms per se can also be used as hosts. Specifically, insect cells such as cells of *Spodoptera frugiperda*, cells of *Bombyx mori*, etc. may be used.

As vectors, there can be used plasmid, phage, phagemid, virus (baculovirus (insect cell expression system), vaccinia virus (animal cell expression system)) and the like.

The vectors of the present invention may contain expression regulatory regions such as a promoter, a terminator, an origin of replication, and the like, depending on the host into which said vector is to be introduced. As promoters for bacterial expression vectors, there can be used commonly used promoters such as trc promoter, tac promoter, lac promoter, and the